U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE

OFFICE NOTE 236

A Relation Between Computer Speed and Errors in Centrally Produced Meteorological Guidance Products

Frederick G. Shuman
National Meteorological Center

FEBRUARY 1981
(Revised JUNE 1981)

This is an unreviewed manuscript, primarily intended for informal exchange of information among NMC staff members.

A Relation between computer speed and errors in centrally produced meteorological guidance products

Frederick G. Shuman
National Meteorological Center

Improvements in central guidance have improved public forecasts substantially, sometimes even dramatically. Figure 1 shows an example from the record. Improvements in central guidance, in turn, have, and still do, depend on a myriad of factors, one of which is faster computation. All of the factors, including faster computation, are essential to improved guidance.

During the 1940's, and early 1950's, before the era of modern computers, improvements came from development of the upper air observing network, followed by discovery of relatively simple approximations to the physical laws governing the behavior of the atmosphere. Even the relatively simple approximations could only roughly be calculated manually, and about a 20% drop in error was experienced when the approximations (barotropic model of the atmosphere) were rather precisely calculated on the IBM-704. Modern computing machinery was essential for this accomplishment, as was the earlier developments in observations, communications networks, and theory. But also essential were timely analyses of the data and timely starts of the calculations. These depended on the successful development of automatic analysis systems, and hardware to automatically transfer data from incoming communications lines into computer memory. Further R&D were also essential, for when the approximations as given were precisely calculated, certain devastating, systematic effects (retrogression of planetary waves, spurious anticyclogenesis) were present in the result.

The IBM-704 was acquired by NOAA in 1957, and all of the essential factors came together in 1958. NOAA has successively replaced its central computing machinery with three generations of more powerful systems, the IBM 7090 in 1960, the CDC 6600 in 1965, and the IBM 360/195 in 1973. Each of these represented at the time more advanced computer technology in both hardware and software. Each generation was accompanied by new developments in modelling the physical laws of the atmosphere, and during the entire period to the present, improvements and new developments in observations (satellites, aircraft, buoys, automatic reduction systems), analysis (influence functions, Hough functions, optimum interpolation), communications (GTS, AWN, automata), and numerical techniques (stable finite-differences, spectral, fast Fourier transform).

With so many and various factors, all more or less continually improving with time, and all to a considerable extent driving each other, one might expect to find a fairly consistent simple relationship between the resultant error in central guidance, and the one factor which has a rather consistent historical variation with time--computer speed. The consistent variation of computer speed with time is shown in Figure 2.

Figure 3, on the other hand, shows that error varies rather consistently with time also. Because both speed and error vary rather simply and consistently with time, the functional relationship that must exist between error and computer speed should be rather simple and consistent.

I would expect a priori the error in central guidance to drop exponentially with some function of computer speed. What function? If a simple relationship exists, I would expect the function to approximate the cube root of computer speed, for the most effective thing done on the IBM 360/195 was increase of horizontal resolution, and the number of calculations increases as the cube of the resolution. I would expect error to vary exponentially with the resolution. Thus I examined the fit of

$$\text{error} = ab^{f^c}$$

where I expected c to be fairly close to 1/3.

As an example, consider the error data for the 500 mb 36 hr guidance product (Fig. 3):

Machine	Speed (MIPS)	f (Speed relative to 195)	Error
(subjective)	0	0	.61
IBM 704	.047	1/225	.48
IBM 7094	.30	1/40	.44
CDC 6600	2	1/6	.32
IBM 360/195	12	1	.223

TABLE I. Values in the error column correspond to those error levels of the bars in Figure 3 which are labelled with machine numbers. The models indicated were designed for the machine as labelled, although in some cases a model designed for one machine was run on its successor for some time. Note also that in the case of the CDC 6600, the error for "Winds input" was used, this being the best of the software applications systems designed for that machinery. This clearly reflects the effect of R&D on error.

The procedure was first to fit the "subjective" data. Thus a = 0.61. Then the IBM 360/195 data were fit: b = 0.223 /0.61 = 0.366. I then found that c = .275 gives a close fit to the other three pairs of data. Thus, I have

$$\text{error} = 0.61 \times 0.366^{f^{0.275}}$$

Figure 4 shows how closely this curve fits all the data.

With such a close fit, I should be able to extrapolate with a high probability, the effect on central guidance of all of the ongoing and future activities to improve models, analyses, observations, communications, and numerical techniques. Much of the total activity depends on computer speed, and I also show in Figure 4 the expected error for $f = 5, 10, 25, 125$, and ∞, where f is the ratio of speed to speed of the IBM 360/195. As in all predictions, the error of extrapolation grows with range from the known data, but since the known data include about three orders of magnitude of f, the extrapolation should well fit future experience out to $f = 10, 20$, or so.

The error in run-of-the-mill cases is already dropping slowly. Much of the future drop in error will reflect greater skill in predicting unusual events--those threatening safety and health, and property.

2/18/81

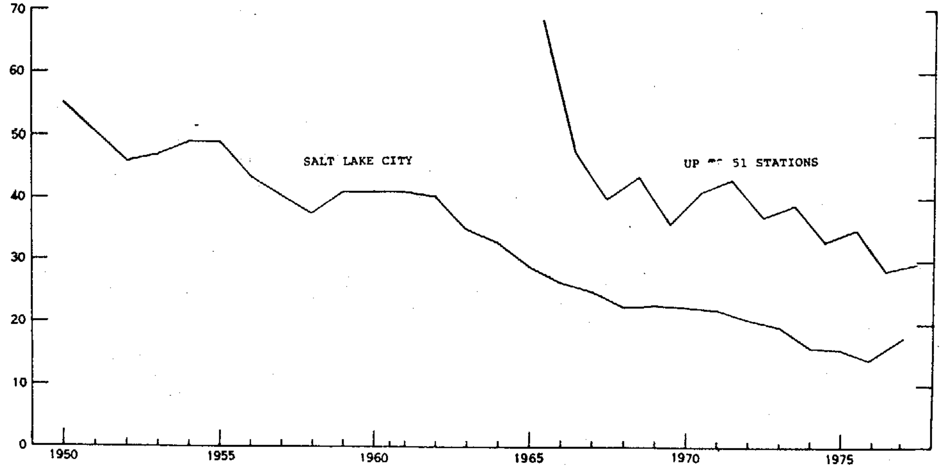

FIG. 1. Forecast temperature errors greater than 10°F. Salt Lake City is the only station that has kept a consistent record of this for over 20 years, and its record is shown to illustrate improvements over a longer period. The shorter record is for forecasts local to WSFOs, which are scattered fairly uniformly about the country, and is the average number per WSFO. Both curves are for forecasts 36 to 48 hr in advance. The Salt Lake City record is for one forecast per day (365 each year). The shorter record is for two forecasts per day, but only for the colder half of the year, October 1 to March 31. The difference in level of the two curves is largely due to two factors: (1) it is more difficult to forecast temperature in the colder seasons because temperature variations are larger, and (2) the variation of temperature over the Great Basin is smaller than elsewhere; for example, over the Great Plains. To remove large fluctuations that appear year-to-year in single-station records, data for Salt Lake City have been smoothed with weighting factors 1:2:4:2:1.

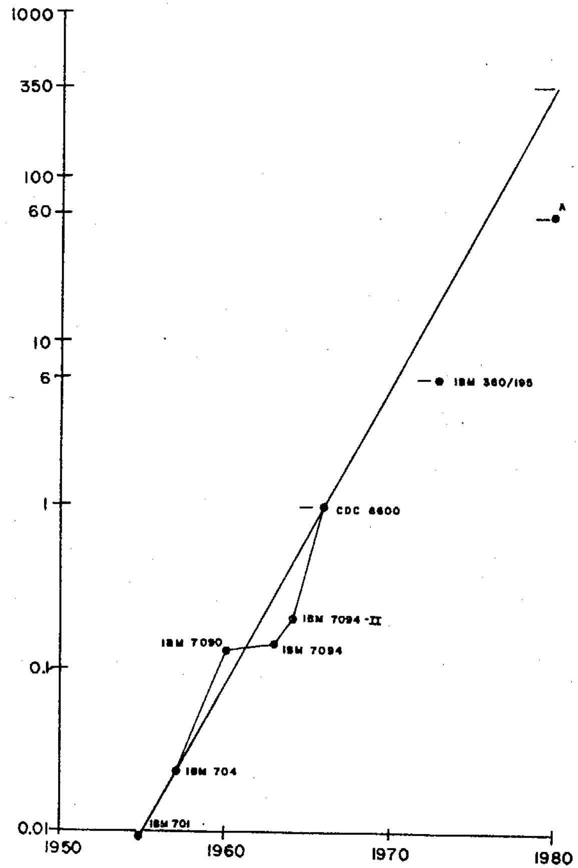

FIG. 2. Operational computer speed, relative to the CDC-6600, plotted against year of acquisition for NMC use.

2/18/81

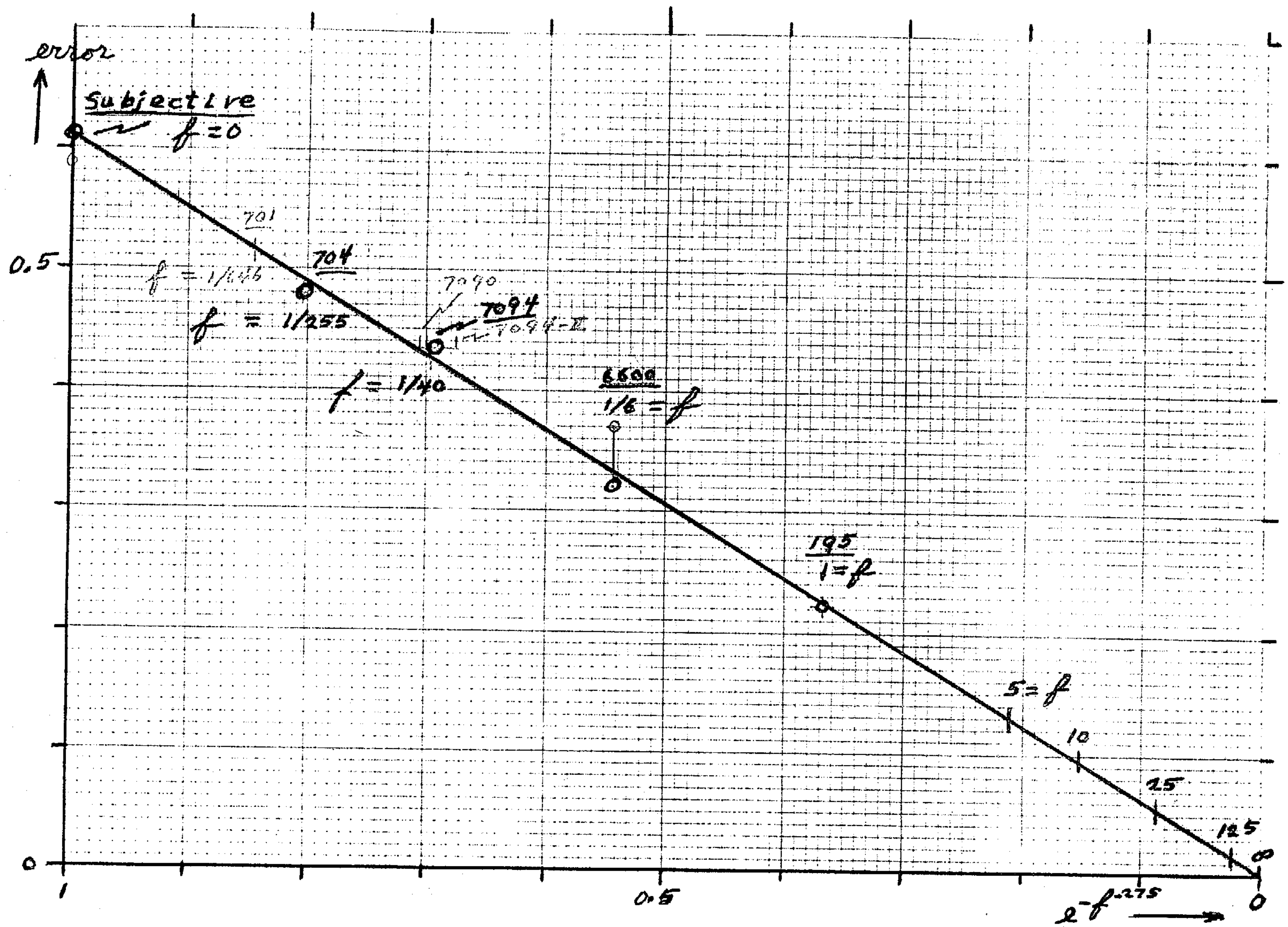

Fig. 4. Error of the NMC 36 hr 500 mb guidance product plotted against a function of computer speed, $e^{-f^{.275}}$. (Note that $e^{-1} \cong 0.366$). f is the ratio of computer speed to the speed of the IBM 360/195. The five small circles are plotted actual data. The straight line is the curve, error = .61 x $e^{-f^{.275}}$, fitted to the data. Error = 1 represents a worthless product; error = 0, a virtually perfect one.

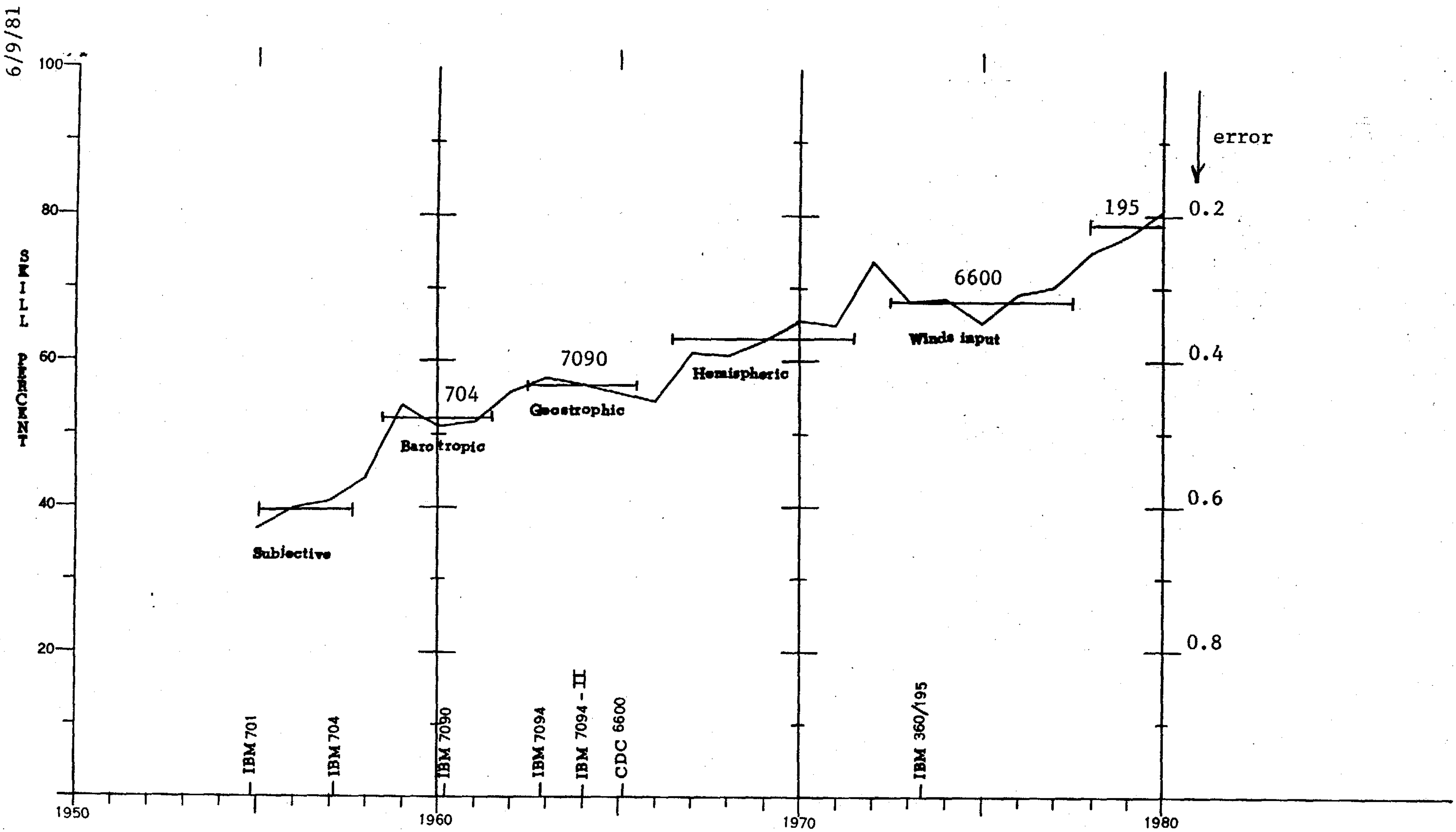

Fig 3. Record of skill, averaged annually, of the NMC 36 h, 500 mb (~ 5.6 km high) circulation predictions over North America. The horizontal bars show averages for the years during which no major changes in models occurred. Years of transition are not included in the averages. The Geostrophic Model (Cressman, 1963) is a generalization of the Barotropic Model, to account for baroclinic effects. The bar labelled "Winds input" shows the effect on skill of a change in input wind. Prior to 1972 a quasi-geostrophic wind field, derived from the pressure field, was used for initial winds. During 1972, analyses of observed winds were used instead. The measure of skill is based on the so-called S_1 score (Teweles and Wobus, 1954), which is a measure of normalized error in horizontal pressure gradients. A chart with an S_1 score of 20 is virtually perfect, and one with 70 is worthless. As shown, skill (percent) is $2 \times (70 - S_1)$, which yields 0 for a worthless chart, and 100 for a virtually perfect one.